Description of the SF-3 Occupation Table Crosswalk, 1990-2000: United States

Introduction

The Census Bureau's website American FactFinder (AFF) now provides data from both the 1990 Census and Census 2000. The data sets used to display industry and occupation (I&O) data on the AFF are the 1990 Summary Tape File 3 (STF-3) and 2000 Summary File 3 (SF-3). The SF-3 on the AFF shows the Census 2000 I&O data in "Detailed Tables," "Quick Tables" and a "Demographic Profile of Economic Characteristics." The Profile provides a subset of the categories found in the detailed tables.

For industry and occupation data, however, you cannot compare the categories directly across the two census years. The wording of the categories is different, and, even when the words *appear* to be the same, the definitions of the categories are sometimes different. For example, "Managers" and "Farming occupations" were defined differently in 1990 than in 2000, even though these words appear in tables from both censuses. These differences were caused by the revisions to the Standard Occupational Classification (SOC) and to the North American Industry Classification System (NAICS). The Census 2000 industry and occupation (I&O) categories were derived from these two standard classifications. See the "Background" section below for more details.

This document provides a description of the first in a series of crosswalks that will enable the Census Bureau and its customers to compare published I&O data between the 1990 Census and Census 2000. The crosswalk series will compare the I&O categories at the various levels of detail published in both censuses, ranging from the very aggregated census profiles and SF-3 tables to the most detailed Equal Employment Opportunity (EEO) and Public Use Microdata Samples (PUMS) files. The ultimate and final crosswalk will be a technical paper similar to the ones published after the 1980 and earlier censuses. There was no such technical paper after the 1990 Census.

The first crosswalk provides the most aggregated (or least detailed) comparisons of the occupation categories only. This crosswalk is "preliminary," and may be revised in the future as the Census Bureau continues its research and as more detailed crosswalks become available. Nevertheless, data users have already made it clear that they prefer having as soon as possible preliminary methods for comparing census I&O data over time, rather than no methods at all. For this reason, the Census Bureau is publishing this first crosswalk to be used in conjunction with the occupation data published in SF-3, including the detailed tables, quick tables, and profiles.

Background

The crosswalks are necessary even more than in previous censuses, because for the first time both the standard industry and occupation classifications underwent major revisions in the same decade. The 1987 Standard Industrial Classification (SIC) was replaced in 1997 by the North American Industrial Classification System (NAICS), and the 1980 Standard Occupational Classification (SOC) was replaced in 1998 by a completely revamped SOC. The 1998 SOC then evolved into a slightly modified update in 2000. The 1997 NAICS and 2000 SOC, respectively, provided the structure for the Census 2000 Industry and Occupation Classifications. The major differences between the 1997 NAICS and 2000 SOC and their previous counterparts became equally major differences between the 1990 Census and Census 2000 classifications. When data users look at census data over time, therefore, they are faced with apparent changes in the industrial and

occupational distributions of the workforce that are due more to classification revisions than to real changes in the workforce itself. Data users need ways to sort out the classification revisions from the real changes. The crosswalks will provide them with the means to do so.

There are two basic methods useful for creating crosswalks. The first is to "double-code" samples of census data into both the 1990 and 2000 I&O classifications. This process in effect puts each person in the sample into both classification sets so that we can examine the flows from each I&O code in one classification, e.g., 1990, into each code in the other classification, e.g., 2000. Another helpful way to look at these flows is to examine the I&O "titles" found in the *Alphabetical and Classified Indexes of Industries and Occupations* published for both censuses.

The matching of I&O titles provides information about the intended, or "ideal," changes from each I&O code of one classification into each I&O code of the other classification. Coding error is minimal. However, this matching does not provide data on the distribution of actual *people* being moved from one category to another.

The double-coding of sample records, on the other hand, provides information on the proportion of actual workers that went from one census category into another. But unfortunately, any coding process involves coding error. In both censuses, these errors average about 7-8 percent for detailed industry codes, and 10-12 percent for occupation codes. These errors contaminate the comparisons across classifications, because they create false combinations of 1990 and 2000 codes.

Given that both the double-coded samples and the index titles methods have advantages and disadvantages by themselves, therefore, this and future crosswalks attempt to merge the best of both methods. With the help of staff at the Bureau of Labor Statistics, the Census Bureau used information from both sources to calculate adjustment factors, apply those factors to census data from the 1990 Census categories, make those categories comparable to the Census 2000 categories, and redistribute the 1990 data based on these factors.

Ideally, different crosswalks could be created based on many different variables, including geography, sex, and race. For example, changes in industrial and occupational distributions are different in New York compared to Kansas, and are different for men and women. To create many different crosswalks depending on all characteristics, however, would require a very large sample controlled for all these variables. Neither financial nor human resources are available to create and analyze such a large sample. For this reason, all the crosswalks produced directly by the methods described above will be national in scope.

Nevertheless, data users may want to do similar comparisons for sub-national geographic areas, such as states, counties, etc. For this reason, the Census Bureau is also providing a "Template," which contains the conversion factors used to create the U.S. occupation crosswalk. Using this Template, data users can apply national conversion factors to any 1990 occupation data. Data users, however, will have to apply the same adjustment factors and distributions to all levels of geography, both sexes, and all other characteristics, even though we know this is not a perfect method for comparing I&O data in all circumstances.

The Data

The entire Census 2000 Occupational classification system consists of 509 detailed civilian and military occupation categories. In SF-3, however, only 33 unique civilian occupational groups are shown, along with 13 more aggregations of those groups (46 table lines altogether). Also published on the AFF are "Profiles of Economic Characteristics," which display only six even more aggregated groupings. The first occupation crosswalk compares the data for all these categories across time by converting the 1990 data into the Census 2000 definitions.

It is also possible to convert the 2000 categories into those for 1990 in order to make comparisons over time. At this time we recommend against this conversion for two reasons: 1) the conversion factors used in these crosswalks were based on a sample of cases from the 1990 Census that was re-coded into the 2000 classifications, and 2) intuitively we prefer to move forward to the present and express the labor force of 1990 in 2000 terms, since Census 2000 is more up-to-date.

The following are highlights of the comparisons between 1990 and 2000:

1. The number of people employed increased by 12.1 percent from 115.7 million in 1990 to 129.7 million in 2000. Occupations showing a percent increase significantly larger than 12.1 percent, therefore, grew at a faster rate than average.

2. By far the occupation group with the largest percent increase was Computer and mathematical occupations, which more than doubled from 1990 to 2000 (113.4 percent increase).

3. The next highest percentage increases, each going up by a third or more from 1990 to 2000, were: Health technologists and technicians (45.2 percent); Supervisors, transportation and material moving occupations (43.7 percent); Community and social service occupations (43.4 percent); Personal care and service occupations (38.4 percent); Fire fighting, prevention, and law enforcement workers, including supervisors (35.4 percent); and Legal occupations (33.6 percent).[1]

4. The categories showing the greatest losses from 1990 to 2000 were: Drafters, engineering, and mapping technicians (-40.0 percent); Extraction (mining) workers (-24.1 percent); Rail, water, and other transportation occupations (-23.4 percent), and Farmers and farm managers (-22.2 percent).

Future Crosswalks

The next crosswalk will be a national-level industry counterpart to the SF-3/Profile Occupation Crosswalk, including a template for creating sub-national crosswalks.

Depending on resources, following the SF-3 level crosswalks, the Census Bureau plans more detailed versions that can be used in conjunction with Census 2000 data products such as SF-4, the EEO File, and the PUMS files. As the more detailed crosswalks are produced, these versions may

[1] Increases and decreases may not be statistically significantly different from each other.

bring about revisions to the earlier crosswalks such as the SF-3 versions being published at this time. The final and most detailed product will be a Technical Paper similar to those published after the 1970 and 1980 Censuses.